Statistical Mechanical Study on a Neural Network Model with Time Dependent Interactions

T Uezu[1], K Abe[1], S Miyoshi[2] and M Okada[3,4]

[1] Graduate School of Sciences and Humanities, Nara Women's University, Nara 630-8506

[2] Department of Electrical and Electronic Engineering, Faculty of Engineering Science, Kansai University, Osaka, 564-8680

[3] Devision of Transdisciplinary Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, 277-8561

[4] RIKEN Brain Science Institute, Saitama, 351-0198

E-mail: uezu@ki-rin.phys.nara-wu.ac.jp

Abstract. We study a neural network model in which both neurons and synaptic interactions evolve in time simultaneously. The time evolution of synaptic interactions is described by a Langevin equation including a Hebbian learning term, and a bias term which is the interactions of the Hopfield model. We assume that synaptic interactions change much slower than neurons and study the stationary states of synaptic interactions by the replica method. We find that the order of the phase transition changes from the second to the first and that the existence regions of the Hopfield attractor and mixed states increase as the coefficient of the learning term increases. We also study the AT stability of solutions and find that the temperature region in which the Hopfield attractor is stable increases as the learning coefficient increases. Theoretical results are confirmed by the direct numerical integration of the Langevin equation. Further, we study the characteristics of the resultant synaptic interactions by partial annealing and find that the stability of the attractor which emerges after partial annealing is enhanced and those of the coexistent attractors are reduced.

PACS numbers: 87.10.-e, 05.20.-y, 84.35.+i

1. Introduction

In this paper, we consider the double dynamics of neurons and synaptic interactions. In particular, we study the case that the time evolution of neurons is so rapid that synaptic interactions are considered to be constant when the average of correlations of neurons are calculated. We study the stationary states of the system by the replica method. In this study, the replica number n can be any value and the $n \to 0$ limit is not taken unlike the usual replica method. The case in which the synaptic interactions do not change we refer to as "quenched", whereas the case in which they change together with neurons we call "annealing". The present case in which synaptic interactions change much slower than neurons is intermediate and is called "partial annealing". Such systems have been studied previously, e.g., the case that synaptic interactions learn the Hebbian rule[1, 2, 3], the case that the replica number is negative[4], and the case that synaptic interactions are divided into a hierarchy of several groups with adiabatically separated and monotonically increasing time-scale[5]. In this paper, we study the effect of the Hebbian learning in partial annealing introducing the coefficient of the Hebbian learning. If the coefficient is negative, we can study the case of unlearning. In this paper, we mainly study the case of learning, and study the states of neurons and the resultant synaptic interactions when the system reaches to the stationary state[6, 7]. Similar analyses are under investigation for the Mexican-hat type interactions[7] and for the Amit model[8].

In the next section, we give the formulation of the model and study the saddle point equations, the AT stability and the phase transitions. In section 3, the results of numerical simulations are presented and are compared with theoretical results. In section 4, we study the nature of interactions generated by partial annealing. In section 5, a summary and discussion are given. In appendix A, we give the analysis of the AT stability compactly.

2. Formulation

2.1. Model

We consider N neurons. Let σ_i represent the state of ith neuron which takes values ± 1, $\sigma_i = 1$ is the firing state and $\sigma_i = -1$ is the rest state. We assume J_{ij} evolves in time according to the following Langevin equation[2]:

$$\tau \frac{d}{dt} J_{ij} = \frac{1}{N} \varepsilon \langle \sigma_i \sigma_j \rangle_{sp} + \frac{1}{N} K_{ij} - \mu J_{ij} + \eta_{ij}(t) \sqrt{\frac{\tau}{N}}, \qquad i < j = 1, \cdots, N \tag{1}$$

The first term $\langle \sigma_i \sigma_j \rangle_{sp}$ is the expectation value of the correlation between ith and jth neurons and represents the Hebbian learning. We assume that this is calculated by the canonical distribution with the Hamiltonian H with the instantaneous values of J_{ij} at

time t.

$$\langle \sigma_i \sigma_j \rangle_{sp} \equiv \frac{1}{Z_\beta} \mathrm{Tr}_{\{\sigma_i\}} e^{-\beta H} \sigma_i \sigma_j, \tag{2}$$

$$H(\{\sigma_i\}) = -\frac{1}{2} \sum_{i \neq j} J_{ij} \sigma_i \sigma_j, \tag{3}$$

$$Z_\beta = \mathrm{Tr}_{\{\sigma\}} e^{-\beta H}. \tag{4}$$

Here, $\mathrm{Tr}_{\{\sigma\}}$ denotes the summation of all configurations of neurons $\{\sigma_i\}$. Z_β is the partition function of neurons. ε is the learning coefficient. The second term K_{ij} is the basic synaptic interactions, to which J_{ij} converges if the learning term and the external noise do not exist. In this paper, as K_{ij} we take the Hopfield model with p patterns,

$$K_{ij} = \frac{K}{\sqrt{p}} \sum_{\nu=1}^{p} \xi_i^\nu \xi_j^\nu. \tag{5}$$

The third term $-\mu J_{ij}$ represents the relaxation and is introduced so that J_{ij} does not diverge. The last term η_{ij} is a white Gaussian random variable with the mean 0 and the following covariance:

$$\langle \eta_{ij}(t) \eta_{kl}(t') \rangle = 2\tilde{T} \delta_{ik} \delta_{jl} \delta(t - t'). \tag{6}$$

Here, $\tilde{T}$ represents the strength of the noise, "noise temperature". The coefficients $\frac{1}{N}$ and $\frac{1}{\sqrt{N}}$ are scaling factors so that the system has non-trivial limit as $N \to \infty$. By defining $\mathcal{H}$ as

$$\mathcal{H} = -\sum_{i<j} K_{ij} J_{ij} + \frac{\mu N}{2} \sum_{i<j} J_{ij}^2 - \frac{\varepsilon}{\beta} \ln Z_\beta, \tag{7}$$

the Langevin equation is rewritten as follows:

$$\tau \frac{dJ_{ij}}{dt} = -\frac{1}{N} \frac{\partial \mathcal{H}}{\partial J_{ij}} + \eta_{ij}(t) \sqrt{\frac{\tau}{N}}. \tag{8}$$

In the stationary state of eq.(8), the probability density $P(\{J_{ij}\})$ of the synaptic interactions $\{J_{ij}\}$ is given by

$$P(\{J_{ij}\}) \propto e^{-\tilde{\beta} \mathcal{H}}, \tag{9}$$

$$\tilde{\beta} = \frac{1}{\tilde{T}}. \tag{10}$$

Thus, the partition function $\tilde{Z}_{\tilde{\beta}}$ of the total system is expressed by

$$\begin{aligned} \tilde{Z}_{\tilde{\beta}} &= \int d\boldsymbol{J} e^{-\tilde{\beta} \mathcal{H}} \\ &= \int d\boldsymbol{J} Z_\beta^n e^{-\frac{\tilde{\beta}}{2} N \mu \sum_{i<j} J_{ij}^2 + \frac{\tilde{\beta} K}{\sqrt{p}} \sum_{i<j} J_{ij} \sum_{\nu=1}^{p} \xi_i^\nu \xi_j^\nu}. \end{aligned} \tag{11}$$

Here, $d\boldsymbol{J} = \prod_{i<j} dJ_{ij}$ and $n = \varepsilon \frac{\tilde{\beta}}{\beta}$. Now, we calculate Z_β^n by the replica method regarding n as an integer. Introducing n replicas $\sigma_i^1, \sigma_i^2, \ldots, \sigma_i^n$, Z_β^n is expressed as

$$\begin{aligned} Z_\beta^n &= \prod_{\alpha=1}^{n} \mathrm{Tr}_{\{\sigma_i^\alpha\}} e^{-\beta H(\{\sigma_i^\alpha\})} \\ &= \mathrm{Tr}_{\{\sigma_i^\alpha\}} e^{\beta \sum_{i<j} J_{ij} \sum_\alpha \sigma_i^\alpha \sigma_j^\alpha + \beta \sum_{\nu,i} h_\nu \xi_i^\nu \sum_\alpha \sigma_i^\alpha}. \end{aligned} \tag{12}$$

We define order parameters m_ν^α and $q^{\alpha\beta}$ as

$$m_\nu^\alpha = \frac{1}{N}\sum_i \sigma_i^\alpha \xi_i^\nu, \tag{13}$$

$$q_{\alpha\beta} = \frac{1}{N}\sum_i \sigma_i^\alpha \sigma_i^\beta. \tag{14}$$

Then, we obtain the following expressions:

$$\tilde{Z}_{\tilde{\beta}} = \sqrt{\frac{2\pi}{\tilde{\beta}}N\mu}\int \frac{iN}{2\pi}d\hat{q}_{\alpha\beta}\int \frac{N}{2\pi i}d\hat{m}_\nu^\alpha e^{NG}, \tag{15}$$

$$G = G_1 + G_2 + G_3, \tag{16}$$

$$G_1 = \frac{1}{4\mu}\tilde{\beta}K^2 + \frac{\beta^2}{2\mu\tilde{\beta}}\sum_{\alpha<\beta} q_{\alpha\beta}^2 + \frac{\beta^2 n}{4\mu\tilde{\beta}} + \frac{\beta K}{2\mu\sqrt{p}}\sum_{\alpha,\nu}(m_\nu^\alpha)^2 + \beta\sum_{\alpha,\nu} h_\nu m_\nu^\alpha, \tag{17}$$

$$G_2 = -\sum_{\alpha<\beta}\hat{q}_{\alpha\beta}q_{\alpha\beta} + \sum_{\alpha,\nu}\hat{m}_\nu^\alpha m_\nu^\alpha, \tag{18}$$

$$G_3 = \frac{1}{N}\sum_i \ln\{\mathrm{Tr}_{\{\sigma_i^\alpha\}} e^{\sum_{\alpha<\beta}\hat{q}_{\alpha\beta}\sigma_i^\alpha\sigma_i^\beta - \sum_{\alpha,\nu}\hat{m}_\nu^\alpha\sigma_i^\alpha\xi_i^\nu}\} \tag{19}$$

$$= \Big[\ln\{\mathrm{Tr}_{\{\sigma^\alpha\}} e^{\sum_{\alpha<\beta}\hat{q}_{\alpha\beta}\sigma^\alpha\sigma^\beta - \sum_{\alpha,\nu}\hat{m}_\nu^\alpha\sigma^\alpha\xi^\nu}\}\Big]. \tag{20}$$

Since we consider the finite number of patterns p and take $N\to\infty$ limit in this paper, $N \gg 2^p$ holds. Therefore, in the first equation of the expression G_3, $\frac{1}{N}\sum_i$ can be replaced by the average over $\{\xi\}$. We denote this average by $[\cdots]$ and get the second equality.

Now, we assume the replica symmetry (RS) as

$$\begin{aligned} q_{\alpha\beta} &= q, \\ m_\nu^\alpha &= m_\nu, \\ \hat{q}_{\alpha\beta} &= \hat{q}, \\ \hat{m}_\nu^\alpha &= \hat{m}_\nu. \end{aligned} \tag{21}$$

Let $G_{i,RS}$ be G_i evaluated at the RS solution. Then, we obtain the following expressions for $G_{1,RS}, G_{2,RS}$ and $G_{3,RS}$:

$$\begin{aligned} G_{1,RS} = {} & \frac{1}{4\mu}\tilde{\beta}K^2 + \frac{\beta^2}{\mu\tilde{\beta}}\frac{n(n-1)}{4}q^2 + \frac{\beta^2 n}{4\mu\tilde{\beta}} + \frac{\beta K}{2\mu\sqrt{p}}n\sum_\nu m_\nu^2 \\ & + \beta n\sum_\nu h_\nu m_\nu, \end{aligned} \tag{22}$$

$$G_{2,RS} = -\frac{n(n-1)}{2}\hat{q}q + n\sum_\nu \hat{m}_\nu m_\nu, \tag{23}$$

$$G_{3,RS} = -\hat{q}\frac{1}{2}n + \Big[\ln\int Dx\Big\{\cosh\Big(\sqrt{\hat{q}}x - \sum_\nu \hat{m}_\nu\xi^\nu\Big)\Big\}^n\Big] + n\ln 2, \tag{24}$$

$$Dx = \frac{dx}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

2.2. Saddle Point Equations (SPEs)

Defining $G_{RS} = G_{1,RS} + G_{2,RS} + G_{3,RS}$, we obtain the following saddle point equations. Here, we put $h_\nu = 0$ and define $J = \dfrac{K}{\mu\sqrt{p}}$ and $\kappa = \dfrac{\beta^2}{\mu\tilde{\beta}}$.

$$\frac{\partial G_{RS}}{\partial q} = 0 : \hat{q} = \kappa q, \tag{25}$$

$$\frac{\partial G_{RS}}{\partial \hat{q}} = 0 : q = \Big[\int Dx \cosh^n \Xi \tanh^2 \Xi \Big\{ \int Dx \cosh^n \Xi \Big\}^{-1} \Big], \tag{26}$$

$$\frac{\partial G_{RS}}{\partial m_\nu} = 0 : \hat{m}_\nu = -\beta J m_\nu, \tag{27}$$

$$\frac{\partial G_{RS}}{\partial \hat{m}_\nu} = 0 : m_\nu = \Big[\xi^\nu \int Dx \cosh^n \Xi \tanh \Xi \Big\{ \int Dx \cosh^n \Xi \Big\}^{-1} \Big], \tag{28}$$

$$\Xi = \sqrt{\kappa q} x + \beta \sum_\nu J m_\nu \xi^\nu. \tag{29}$$

In this paper, we study the case of $p = 3$ and abbreviate the SPEs as

$$q \equiv \varphi(q, m_1, m_2, m_3), \tag{30}$$

$$m_\nu \equiv \psi_\nu(q, m_1, m_2, m_3), \ (\nu = 1, 2, 3). \tag{31}$$

We find the following solutions of the SPEs.

- P : $q = 0$, $m_\mu = 0$, paramagnetic solution,
- SG : $q > 0$, $m_\mu = 0$, spin-glass solution,
- H : $q > 0$, $m_1 \neq 0$, $m_2 = m_3 = 0$, Hopfield attractor,
- $2M$: $q > 0$, $m_1 = m_2 \neq 0$, $m_3 = 0$, mixed state with 2 patterns,
- $3M$: $q > 0$, $m_1 = m_2 = m_3 \neq 0$, mixed state with 3 patterns.

In this paper, we analyze the Hopfield attractor (H), the mixed state with 3 patterns ($3M$) and spin-glass state (SG). We do not consider the mixed state with 2 patterns ($2M$) because it is expected to be unstable.

2.3. AT stability

We study the AT stability of the RS solution[9]. The condition of the stability is that the free energy increases when order parameters deviate from those at the RS solution. The free energy per neuron $\tilde{f}_{\tilde{\beta}}$ is given by

$$\tilde{f}_{\tilde{\beta}} = -\frac{1}{N}\frac{1}{\tilde{\beta}} \ln \tilde{Z}_{\tilde{\beta}} = -\frac{G}{\tilde{\beta}}. \tag{32}$$

We define small deviations from the RS solution by

$$m_\nu^\alpha = m_\nu + \epsilon_\nu^\alpha,$$

$$q_{\alpha\beta} = q + \eta^{\alpha\beta},$$

and expand G up to the second order of deviations. Then, G is expressed as

$$\begin{aligned} G \quad &= G_{RS} + \frac{1}{2} \sum_{(\alpha\nu)(\beta\mu)} \mathcal{G}_{(\alpha\nu)(\beta\mu)} \epsilon_\nu^\alpha \epsilon_\mu^\beta + \frac{1}{2} \sum_{(\alpha\nu)(\beta\gamma)} \mathcal{G}_{(\alpha\nu)(\beta\gamma)} \epsilon_\nu^\alpha \eta^{\beta\gamma} \\ &\quad + \frac{1}{2} \sum_{(\alpha\beta)(\gamma\delta)} \mathcal{G}_{(\alpha\beta)(\gamma\delta)} \eta^{\alpha\beta} \eta^{\gamma\delta}, \qquad (33) \\ G_{RS} &= G \mid_{q_{\alpha\beta}=q, m_\nu^\alpha = m_\nu} . \end{aligned}$$

$\mathcal{G}$ is called the Hessian. The stability condition of the RS solution is that all eigenvalues of $\mathcal{G}$ are negative. We calculated all eigenvalues for H, $3M$ and SG. There exist 7 different kinds of eigenvalues for H and $3M$, and 5 different kinds of eigenvalues for SG. Details are given in Appendix.

2.4. Phase Transition

In this subsection, we study the phase transition. The second order phase transition temperatures are determined by the following relations:

$$T_{P\to H}^{2nd} = T_{P\to 3M}^{2nd} = \frac{K}{\mu\sqrt{p}} = J, \qquad (34)$$

$$T_{P\to SG}^{2nd} = \frac{1}{\sqrt{\mu\tilde{\beta}}} = \sqrt{\frac{\tilde{T}}{\mu}}, \qquad (35)$$

$$T_{SG\to H}^{2nd} = J\Big\{\Big(\varepsilon \frac{T_{SG\to H}^{2nd}}{\tilde{T}} - 1\Big)q + 1\Big\}, \qquad (36)$$

$$T_{SG\to 3M}^{2nd} = T_{SG\to H}^{2nd}. \qquad (37)$$

Since the second order phase transition temperature from P to H, $T_{P\to H}^{2nd}$, and that from P to $3M$, $T_{P\to 3M}^{2nd}$, are equal, we denote them by $T_{P\to M}^{2nd}$. Here M implies both of H and $3M$. Similarly, since $T_{SG\to H}^{2nd}$ and $T_{SG\to 3M}^{2nd}$ are equal, we denote them by $T_{SG\to M}^{2nd}$. Next, we study the first order phase transition. In this case, a new phase appears suddenly irrespective of the old phase from which the transition takes place. Thus, we consider the following three phase transitions and obtain the equations to determine the phase transition temperatures.

(i) Transition to H

$$q = \varphi(q, m_1, 0, 0), \qquad (38)$$

$$m_1 = \psi_1(q, m_1, 0, 0), \qquad (39)$$

$$\Big(1 - \frac{\partial\varphi}{\partial q}\Big)\Big(1 - \frac{\partial\psi_1}{\partial m_1}\Big) - \frac{\partial\varphi}{\partial m_1}\frac{\partial\psi_1}{\partial q} = 0. \qquad (40)$$

(ii) Transition to $3M$

$$q = \varphi(q, m, m, m), \qquad (41)$$

$$m = \psi_1(q, m, m, m), \qquad (42)$$

$$\Big(1 - \frac{\partial\varphi}{\partial q}\Big)\Big(1 - \frac{\partial\psi_1}{\partial m}\Big) - \frac{\partial\varphi}{\partial m}\frac{\partial\psi_1}{\partial q} = 0. \qquad (43)$$

Derivatives in the above equations are calculated as

$$\begin{aligned}\frac{\partial \psi_1}{\partial m} &= \frac{\partial}{\partial m}\psi_1(q,m,m,m) \\ &= \frac{\partial \psi_1}{\partial m_1}+\frac{\partial \psi_1}{\partial m_2}+\frac{\partial \psi_1}{\partial m_3} = \frac{\partial \psi_1}{\partial m_1}+2\frac{\partial \psi_1}{\partial m_2}, \\ \frac{\partial \varphi}{\partial m} &= \frac{\partial \varphi}{\partial m_1}+\frac{\partial \varphi}{\partial m_2}+\frac{\partial \varphi}{\partial m_3} = 3\frac{\partial \varphi}{\partial m_1}.\end{aligned}$$

(iii) Transition to SG

$$q = \varphi(q,0,0,0), \tag{44}$$

$$\frac{\partial \varphi}{\partial q} = 1. \tag{45}$$

We solve the above equations numerically and determine the phase transition temperatures.

Hereafter, we fix $K = 1$ and $\mu = 1$ and change $T, \tilde{T}$ and ε. We display the

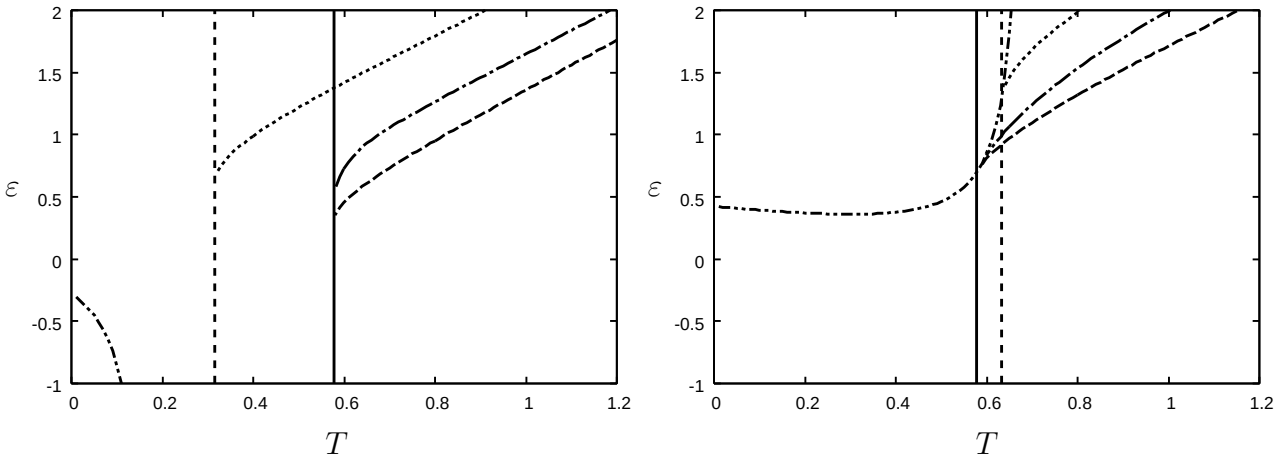

Figure 1. Phase transition lines in $T-\varepsilon$ plane. Left panel: $\tilde{T} = 0.1$. Right panel: $\tilde{T} = 0.4$. $K = 1.0$ and $\mu = 1.0$. Solid curve: $T^{2nd}_{P\to M}$, short dashed curve:$T^{2nd}_{P\to SG}$, dashed-dotted-dotted curve:$T^{2nd}_{SG\to M}$, long dashed curve:T^{1st}_{H}, dashed-dotted curve:T^{1st}_{3M}, dotted curve:T^{1st}_{SG}.

phase transition lines in $T-\varepsilon$ plane for $\tilde{T} = 0.1$ and 0.4 in Fig. 1, and those in $T-\tilde{T}$ plane for $\varepsilon = 1.0$ and $\varepsilon = 3.0$ in Fig. 2, respectively. From these figures, we note the following. There is a tendency that the order of the phase transition changes from the second order to the first one as ε increases for H, $3M$ and SG or $\tilde{T}$ decreases for H and $3M$. The interval of the temperature in which a solution exists increases as ε increases for H, $3M$ and SG or $\tilde{T}$ decreases for H and $3M$. In the right panel of Fig. 3, we display the temperature dependence of q for $\tilde{T} = 0.1$ and $\varepsilon = 1.0$. For three cases of H, $3M$ and SG, when the temperature is decreased, a pair of solutions emerge at the temperature of the first order phase transition, and two solutions are separated

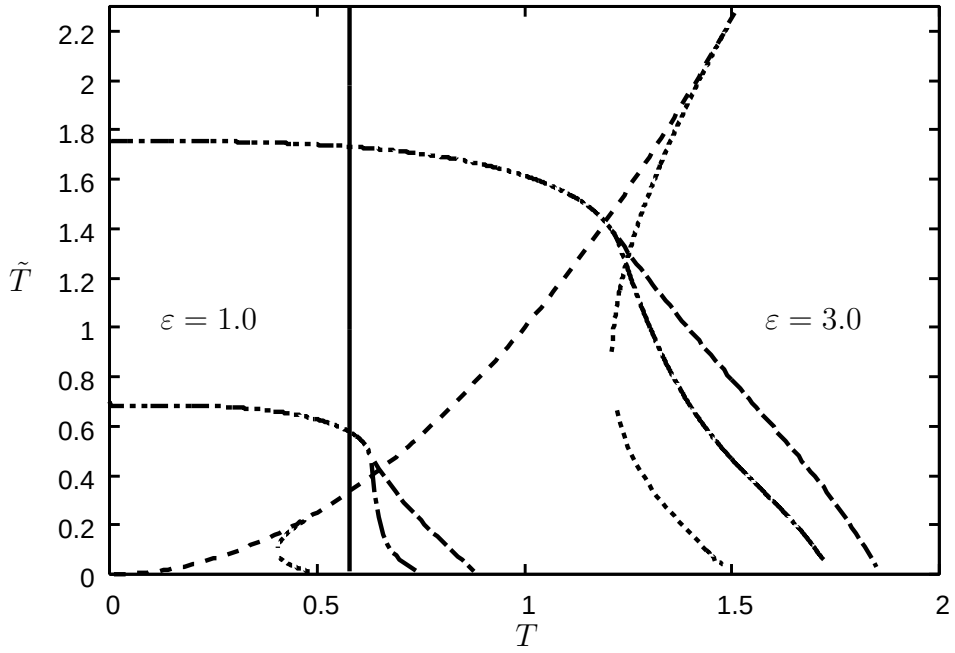

Figure 2. Phase transition lines in $T - \tilde{T}$ plane for $\varepsilon = 1.0$ and $\varepsilon = 3.0$. $K = 1.0$ and $\mu = 1.0$. Solid curve: $T^{2nd}_{P\to M}$, short dashed curve:$T^{2nd}_{P\to SG}$, dashed-dotted-dotted curve:$T^{2nd}_{SG\to M}$, long dashed curve:T^{1st}_{H}, dashed-dotted curve:T^{1st}_{3M}, dotted curve:T^{1st}_{SG}.

and consist of the upper and lower branches as the temperature is further decreased. Examining the AT stability, we find that the upper branch is stable and lower branch is unstable for H and $3M$, whereas the both branches of SG are unstable. As seen

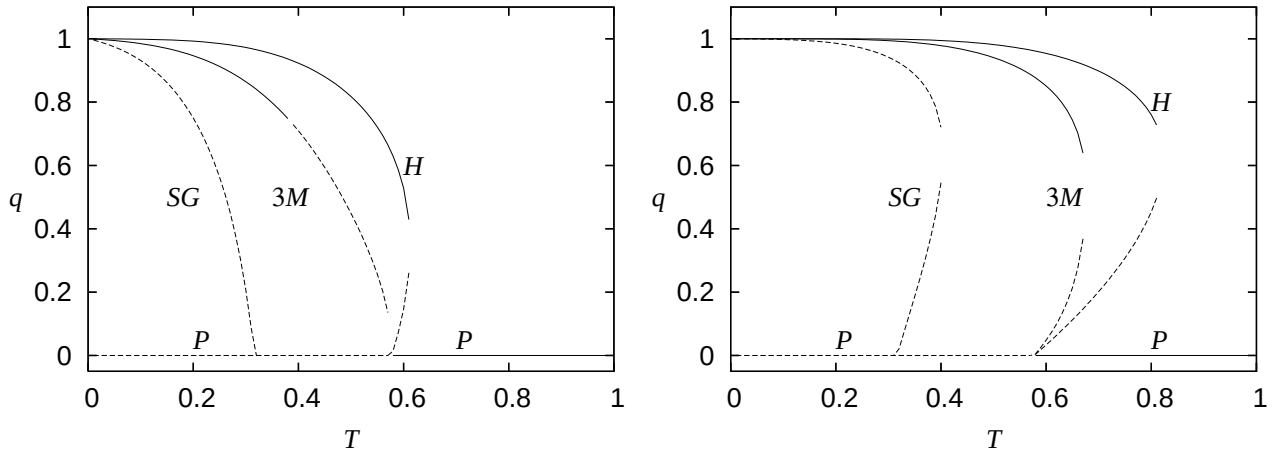

Figure 3. Temperature dependence of q. $\tilde{T} = 0.1$. Left panel: $\varepsilon = 0.5$. Right panel $\varepsilon = 1.0$. $K = 1.0$ and $\mu = 1.0$. Solid curve: stable solution. Dotted curve: unstable solution.

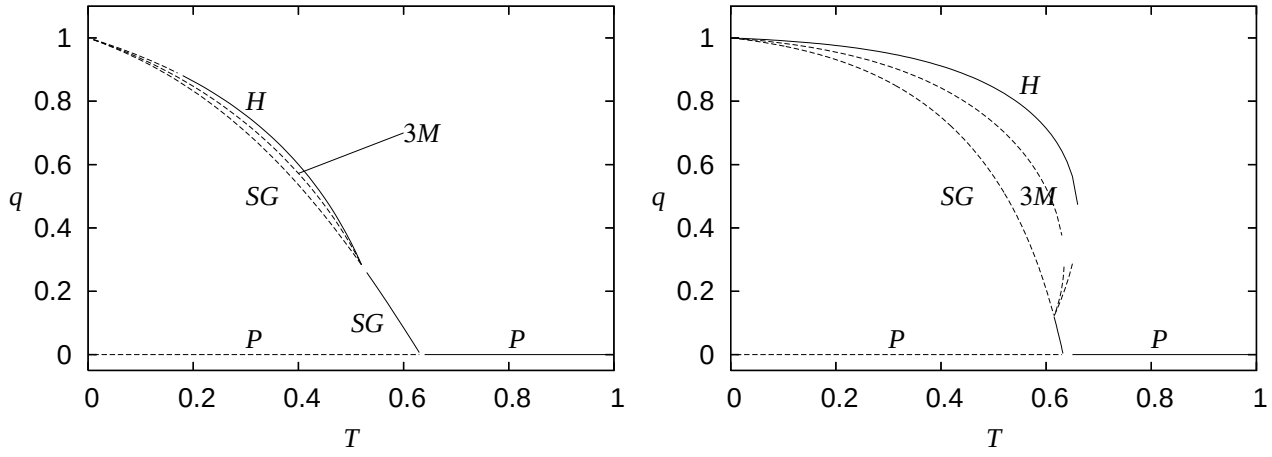

Figure 4. Temperature dependence of q. $\tilde{T} = 0.4$. Left panel: $\varepsilon = 0.5$. Right panel $\varepsilon = 1.0$. $K = 1.0$ and $\mu = 1.0$. Solid curve: stable solution. Dotted curve: unstable solution.

in the left panel of Fig. 3 for $\tilde{T} = 0.1$ and $\varepsilon = 0.5$, H emerges by the first order phase transition, and SG emerges by the second order one. For $3M$, it is difficult to determine the order of the phase transition in this figure, but by Fig. 5 we find that it is the second order. The upper branch of H is stable whenever it exists, but $3M$ is stable only at low temperatures.

In Fig. 4, we display the temperature dependence of q for $\tilde{T} = 0.4$, the left panel is for $\varepsilon = 0.5$ and the right panel is for $\varepsilon = 1.0$. We note that H becomes unstable in low temperatures for $\varepsilon = 0.5$ and that there exist regions where SG is stable.

The second order phase transition temperature $T^{2nd}_{P\to H} = T^{2nd}_{P\to 3M}$ depends on $\tilde{T}$ but not on ε. From this fact together with the fact that when the first order phase transition takes place the upper branches of H and $3M$ are stable for large ε and small $\tilde{T}$ and the upper branch of H is stable for large ε and large $\tilde{T}$, we can understand that the stable temperature region of the upper branch increases when ε is increased with fixed $\tilde{T}$. In Fig. 5, we display the phase diagram in $T - \varepsilon$ space taking into account the stability of the solutions.

In Table 1, we show the phase transition temperature of H, T_c^H, that of $3M$, T_c^M, and the ratio r of the stable temperature region for $3M$ to that for H by taking into account the AT stability for several values of ε and $\tilde{T} = 0.1$. We note that T_c^H, T_c^M and r increases as ε increases from 0.

3. Simulations

We performed direct integrations of the Langevin equation (1) fixing $\tau = 1, K = 1$ and $\mu = 1$ and changing parameters $T, \tilde{T}$ and ε. We used the Euler method with the time increment $\Delta t = 0.1$. We adopted the following procedure according to ref. 2). When

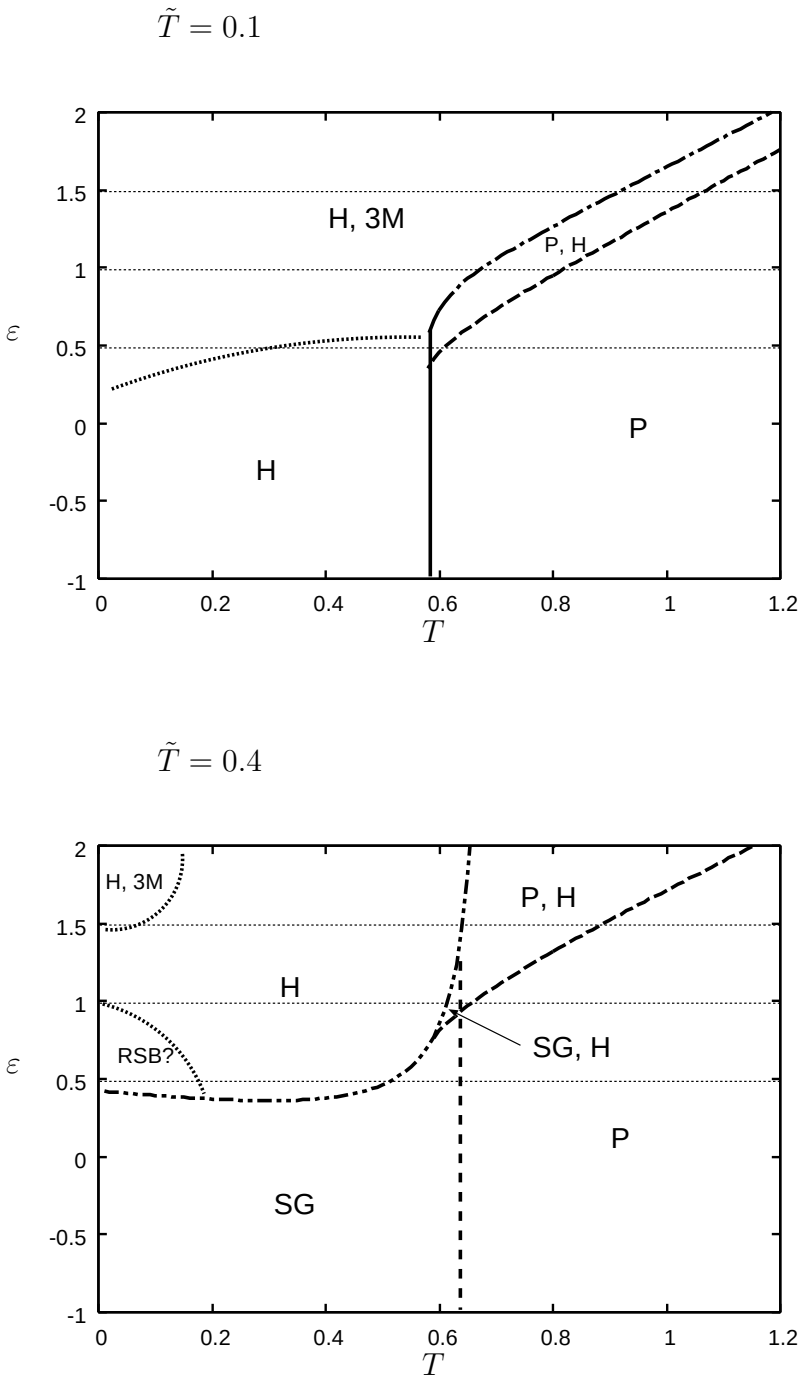

Figure 5. Phase diagram in $T - \varepsilon$ space. $K = 1.0$, $\mu = 1.0$, Upper panel: $\tilde{T} = 0.1$, Lower panel: $\tilde{T} = 0.4$. Solid curve: $T^{2nd}_{P\to M}$, short dashed curve:$T^{2nd}_{P\to SG}$, dashed-dotted-dotted curve:$T^{2nd}_{SG\to M}$, long dashed curve:T^{1st}_{H}, dashed-dotted curve:T^{1st}_{3M}. Dotted horizontal lines indicate the parameter where simulations were performed. Dotted curves denote phase boundaries which are not calculated theoretically but estimated by numerical data available.

Table 1. $\tilde{T} = 0.1, K = 1, \mu = 1.0$. Phase transition temperature of the Hopfield attractor (H) and the mixed state with 3 patterns ($3M$), and the ratio r of the stable temperature region for $3M$ to that for H by taking into account the AT stability for several values of ε.

	H	$3M$	$r = \frac{\text{stable region for } 3M}{\text{stable region for } H}$
$\varepsilon = 0$	0.58	0.27	0.47
$\varepsilon = 0.5$	0.61	0.38	0.62
$\varepsilon = 1.0$	0.83	0.68	0.82
$\varepsilon = 1.5$	1.07	0.92	0.86

N times of update by the Monte Carlo method are tried, we call it 1 Monte Carlo step and denote it by 1[MCS].

(i) Set an initial state of interactions $\boldsymbol{J}(0)$.

(ii) Update neurons $\boldsymbol{\sigma}$ by R_1 [MCS].

(iii) Calculate $\langle \sigma_i \sigma_j \rangle$ during the R_2 [MCS] update of neurons.

(iv) Update $\boldsymbol{J}$ by the Euler method.

(v) Repeat 2 to 4 R_3 times.

(vi) Calculate averages of physical quantities during the R_4 repetitions of 2-4.

In this procedure, the total number of updates of neurons is $(R_1 + R_2)(R_3 + R_4)$ [MCS]. As Rs, we took $R_1 = R_2 = R_3 = R_4 = 500$.

First, we show the numerical results together with theoretical ones for the Hopfield attractor in Fig.6. Parameters are $\tilde{T} = 0.1$, $\varepsilon = 0.5, 1.0, 1.5$ and $N = 1000$. Numerical results agree fairly well with the upper branch solution obtained theoretically. Next, the theoretical and numerical results are shown for the mixed state, in Fig.7. Parameters are $\tilde{T} = 0.1$, $\varepsilon = 0.5, 1.0.1.5$ and $N = 2000$. The agreement between simulations and the stable upper branch solution by theory are quite well, except for two cases. One is just above the critical temperature for $\varepsilon = 1.0$ and $\varepsilon = 1.5$. In the paramagnetic phase, fluctuations are very large. The other is around the temperature where the AT instability takes place for $\varepsilon = 0.5$. In this case, the phase transition takes place at the temperature which is lower than the theoretical prediction. Results in both cases are considered to be finite size effects.

4. On the nature of interactions generated by partial annealing

Here, we numerically study the interaction $\{J_{ij}\}$ which appears when the system reaches to the stationary state. There are two types of such interactions. One is $\{J_{ij}^H\}$ for which H emerges at the stationary state, and the other is $\{J_{ij}^M\}$ for which the $3M$ emerges. Using these two interactions, we performed Monte Carlo simulations taking two initial conditions, H and $3M$.

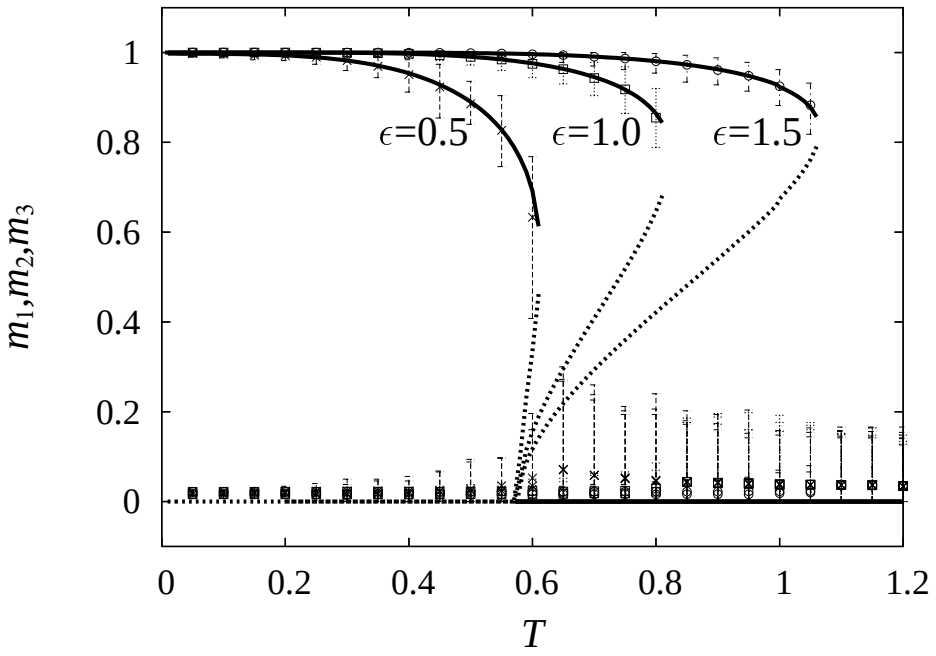

Figure 6. Temperature dependence of m_1, m_2, m_3 for the Hopfield attractor for $\varepsilon = 0.5, \varepsilon = 1.0$ and $\varepsilon = 1.5$. $\tilde{T} = 0.1, K = 1.0$ and $\mu = 1.0$. Curves are theoretical results. Solid curves are stable and dotted curves are unstable. Symbols are simulation results for $N = 1000$.

In Fig. 8, we display the time series of m_μ for $\tilde{T} = 0.1, \varepsilon = 1.0$ and $T = 0.4$. We found that when the interaction is $\{J_{ij}^H\}$, the neuron system converges to H, whereas when the interaction is $\{J_{ij}^M\}$, it converges to $3M$, irrespective of initial conditions. In Fig. 9, we display the theoretical and simulation results for the temperature dependence of m_1, m_2 and m_3 for $\tilde{T} = 0.1$ and $\varepsilon = 1.0$. In the left (right) panel of Fig. 9, at each temperature, the interaction $\{J_{ij}^H\}$ ($\{J_{ij}^M\}$) is used and the initial condition is $3M$ (H). As seen from these figures, the resultant attractor by the Monte Carlo simulation is H for $\{J_{ij}^H\}$ and $3M$ for $\{J_{ij}^M\}$ respectively, irrespective of the initial conditions. This implies that partial annealing has the effect to enhance the stability of the resultant attractor and to reduce that of coexisting attractors.

5. Summary and discussion

We investigated the change of the system behaviors by partial annealing in which the synaptic weights change but much slower than the neurons. As the basic interaction, we took that of the Hopfield model, and introduced the coefficient ε of the Hebbian learning. We studied the stationary states of the Langevin equation by changing parameters, those

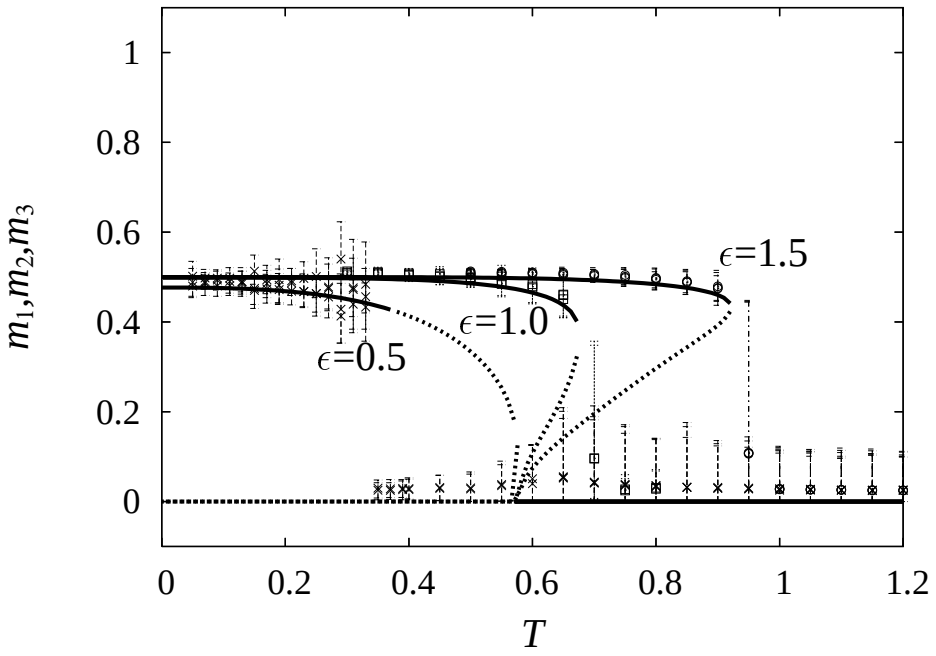

Figure 7. Temperature dependence of m_1, m_2, m_3 for the mixed state 3M for $\varepsilon = 0.5, \varepsilon = 1.0$ and $\varepsilon = 1.5$. $\tilde{T} = 0.1, K = 1.0$ and $\mu = 1.0$. Curves are theoretical results. Solid curves are stable and dotted curves are unstable. Symbols are simulation results for $N = 2000$.

are the learning coefficient ε, the neuron "temperature" T and the synaptic weight "temperature" $\tilde{T}$.

First, we studied the phase transition lines in (T, ε) and $(T, \tilde{T})$ planes and found there is a tendency that the order of phase transitions of solutions becomes the second to the first as ε increases or $\tilde{T}$ decreases.

Next, by taking into account the AT stability, we drew phase diagrams in (T, ε) plane for $\tilde{T} = 0.1$ and 0.4. For $\tilde{T} = 0.1$, we found that the stable parameter regions for the Hopfield attractor and the mixed state with 3 patterns increase as ε increases. Further, we found that the ratio of the stable region of $3M$ to that of the Hopfield attractor increases as ε increases. For $\tilde{T} = 0.4$, we found that SG appears by the second order phase transition from P for small ε and it is the only stable solution. When ε is large, the Hopfield attractor appears by the first order phase transition, and the stable region of the Hopfield attractor increases as ε increases. We confirmed the theoretical results by the numerical simulations, except for the finite size effects observed only for a few set of parameters.

In order to study the nature of interactions generated by partial annealing, we performed the Monte Carlo simulations using the synaptic weights obtained by partial

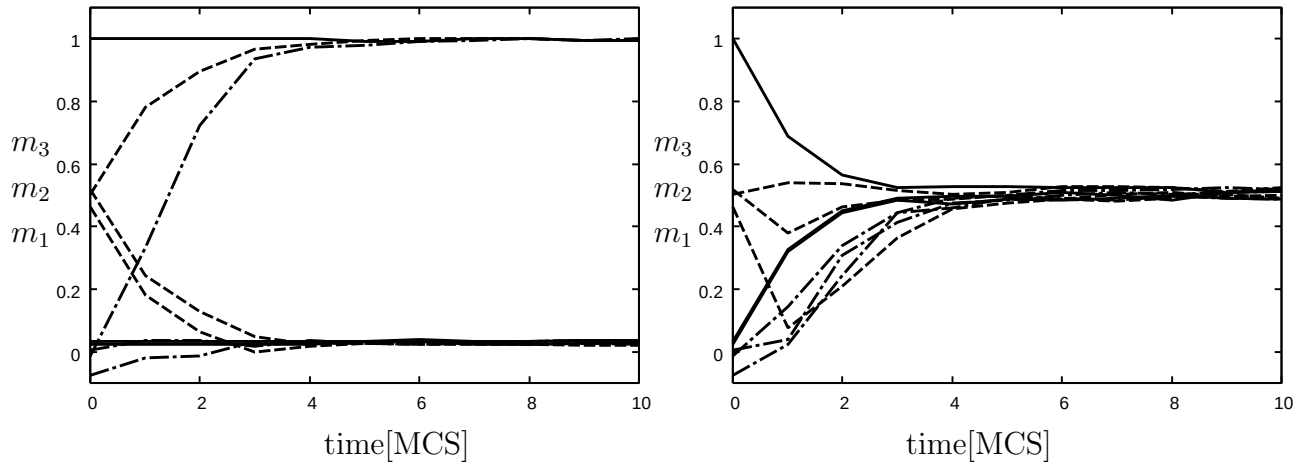

Figure 8. Time series of m_1, m_2, m_3. Numerical results ($N = 500$). $K = 1, \mu = 1, \tau = 1$. $\tilde{T} = 0.1, \epsilon = 1.0, T = 0.4$. Left panel. Interaction $\{J_{ij}^H\}$, in which H appears during partial annealing, is used. Right panel. Interaction $\{J_{ij}^M\}$, in which $3M$ appears during partial annealing, is used. Initial state. Solid curve: $m_1 = 1, m_2 = m_3 = 0$. Dashed curve: $m_1 \simeq m_2 \simeq m_3 \simeq 0.5$. Dashed dotted curve: random configuration.

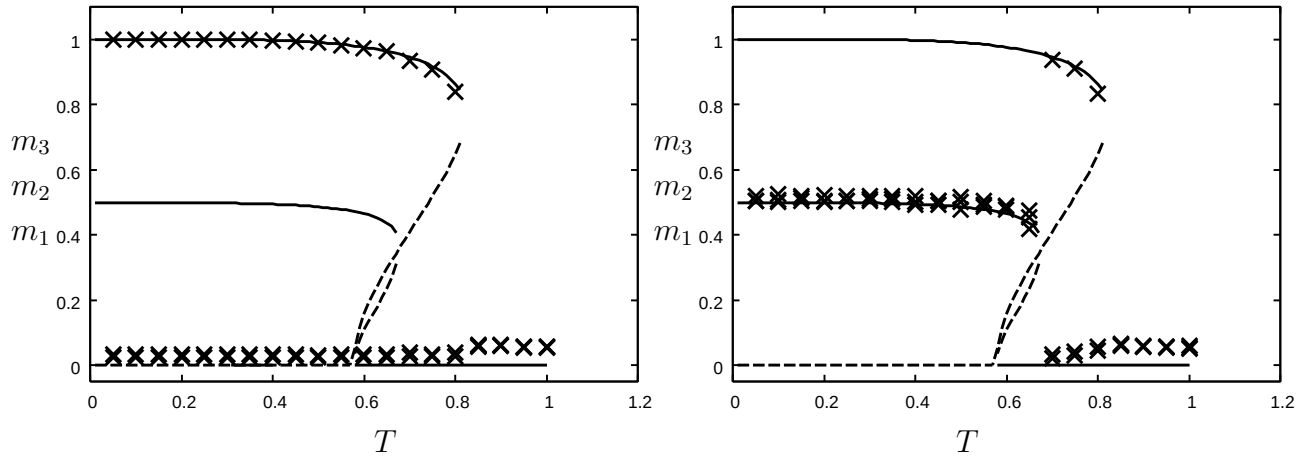

Figure 9. Temperature dependence of m_1, m_2, m_3. $\tilde{T} = 0.1, \epsilon = 1.0$. Curves are theoretical results of partial annealing and symbols are simulation results ($N = 500$). Left panel. At each temperature, interaction $\{J_{ij}^H\}$ in which H appears during partial annealing is used. Initial state is $3M$. Right panel. At each temperature, interaction $\{J_{ij}^M\}$ in which $3M$ appears during partial annealing is used. Initial state is H.

annealing. As a result, we found that the stability of the attractor which emerges as the result of partial annealing is enhanced and that of the coexistent attractors are reduced.

In the present model, the case of $\tilde{T} = 0$ and $\varepsilon = 0$ is nothing but the original Hopfield model. In this case, the Hopfield attractor and $3M$ are stable for low temperatures. We numerically confirmed that the Hopfield attractor is stable for all values of ε when it exists for $\tilde{T} = 0.1$ although there is a possibility that it may be AT unstable when T is very low for $\tilde{T} > 0$. On the other hand, for $\tilde{T} = 0.4$ we found the region in which the Hopfield attractor becomes AT unstable with the replicon mode eigenvalue $\lambda_3 > 0$ indicating replica symmetry breaking. This happens when ε is less than 1. Thus, it is concluded that if ε is larger than some value which depends on $\tilde{T}$, the larger ε is, the wider the temperature region of the stable Hopfield attractor is.

From Fig. 5, we note that for small value of ε, partial annealing can not widen the stability region of the Hopfield attractor. In particular, it seems that when ε becomes negative, i.e. in the case of unlearning, no significant change happens in the phase diagram. We consider that this is because the loading rate of patterns, $\alpha = \frac{p}{N}$, is 0 in this study. If α is positive and large, results might change. This is an interesting unsolved problem.

Appendix A. AT stability of solutions

In this Appendix, the eigenvalues and eigenvectors of the Hessian are calculated. In the below, the symbols α, β, γ and δ indicate the replica indexes and symbols μ and ν do the pattern induces.

Let the Hessian matrix $\mathcal{G}$ be the $L \times L$ matrix. L is the summation of the dimensions of $\{\epsilon_\nu^\alpha\}$ space and that of $\{\eta^{\alpha\beta}\}$ space, i.e., $L = 3n + {}_nC_2 = \dfrac{n(n+5)}{2}$ for $p = 3$. The eigenvalue equation is expressed as

$$\mathcal{G}\boldsymbol{\mu} = \lambda\boldsymbol{\mu}, \tag{A.1}$$

where λ is an eigenvalue of $\mathcal{G}$ and $\boldsymbol{\mu}$ is the eigenvector belonging to λ. $\boldsymbol{\mu}$ takes the following form:

$$\boldsymbol{\mu} = \begin{pmatrix} \{\epsilon_\nu^\alpha\} \\ \{\eta^{\alpha\beta}\} \end{pmatrix}. \tag{A.2}$$

$\{\epsilon_\nu^\alpha\}$ and $\{\eta^{\alpha\beta}\}$ denote a $3n$ dimensional and a ${}_nC_2$ dimensional column vectors, respectively. In the below, $[\cdots]$ is the average over patterns ξ and a bar denotes the following average:

$$\overline{f(\Xi)} \equiv \Omega^{-1} \int Dx \cosh^n(\Xi) f(\Xi), \tag{A.3}$$

$$\Omega \equiv \int Dx \cosh^n(\Xi), \tag{A.4}$$

$$\Xi = \beta\Big\{\sqrt{\frac{q}{\mu\tilde{\beta}}}x + \frac{K}{\mu\sqrt{p}}\sum_\nu m_\nu\xi^\nu\Big\} = \sqrt{\kappa q}x + \beta J\sum_\nu m_\nu\xi^\nu, \tag{A.5}$$

$$\kappa \equiv \frac{\beta^2}{\mu\tilde{\beta}},\ J \equiv \frac{K}{\mu\sqrt{p}}. \tag{A.6}$$

Further, $\langle\cdots\rangle$ denotes the following average at the replica symmetric solution.

$$\begin{aligned}\langle\cdots\rangle &= \frac{\mathrm{Tr}_{\boldsymbol{\sigma}} e^{\tilde{H}}\cdots}{\mathrm{Tr}_{\boldsymbol{\sigma}} e^{\tilde{H}}},\\ \tilde{H} &= \kappa\sum_{\alpha<\beta}(q_{\alpha\beta})^2\sigma^\alpha\sigma^\beta + \beta J\sum_{\alpha\nu} m_\nu^\alpha\sigma^\alpha\xi^\nu\\ &= \kappa q^2\sum_{\alpha<\beta}\sigma^\alpha\sigma^\beta + \beta J\sum_{\alpha\nu} m_\nu\sigma^\alpha\xi^\nu.\end{aligned}$$

Appendix A.1. Hopfield attractor

We put $m_1 = m, m_2 = m_3 = 0$. Then, the non-zero elements of $\mathcal{G}$ are the 7 quantities defined as

$$\mathcal{G}_{(\alpha\nu)(\alpha\nu)} \equiv \frac{\partial^2 G}{\partial m_\nu^{\alpha 2}} = -\beta J(1-\beta J(1-m^2)) = A, \tag{A.7}$$

$$\mathcal{G}_{(\alpha\nu)(\beta\nu)} \equiv \frac{\partial^2 G}{\partial m_\nu^\alpha \partial m_\nu^\beta} = (\beta J)^2(q-m^2) = B, \qquad (\alpha\neq\beta), \tag{A.8}$$

$$\begin{aligned}\mathcal{G}_{(\alpha\beta)(\alpha\beta)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta}\partial q_{\alpha\beta}} &= -\kappa[1-\kappa(1-\langle\sigma^\alpha\sigma^\beta\rangle^2)]\\ &= -\kappa[1-\kappa(1-q^2)] = P, \qquad (\alpha\neq\beta).\end{aligned} \tag{A.9}$$

$$\begin{aligned}\mathcal{G}_{(\alpha\beta)(\alpha\gamma)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta}\partial q_{\alpha\gamma}} &= \kappa^2[\langle\sigma^\beta\sigma^\gamma\rangle - \langle\sigma^\alpha\sigma^\beta\rangle^2]\\ &= \kappa^2(q-q^2) = Q,\ \alpha,\beta \text{ and } \gamma \text{ are all different.}\end{aligned} \tag{A.10}$$

$$\begin{aligned}\mathcal{G}_{(\alpha\beta)(\gamma\delta)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta}\partial q_{\gamma\delta}} &= \kappa^2[\langle\sigma^\alpha\sigma^\beta\sigma^\gamma\sigma^\delta\rangle - \langle\sigma^\alpha\sigma^\beta\rangle^2] = \kappa^2(\overline{\tanh^4\Xi_1} - q^2) = R,\\ &\qquad \alpha,\beta,\gamma \text{ and } \delta \text{ are all different.}\end{aligned} \tag{A.11}$$

$$\mathcal{G}_{(\alpha\beta)(\alpha\nu)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta}\partial m_\nu^\alpha} = \kappa\beta Jm(1-q)\delta_{\nu 1} = C\delta_{\nu 1}, (\alpha\neq\beta). \tag{A.12}$$

$$\mathcal{G}_{(\beta\gamma)(\alpha\nu)} \equiv \frac{\partial^2 G}{\partial q_{\beta\gamma}\partial m_\nu^\alpha} = \kappa\beta J(\overline{\tanh^3\Xi_1} - qm)\delta_{\nu 1} = D\delta_{\nu 1}, \qquad (\alpha\neq\beta, \alpha\neq\gamma). \tag{A.13}$$

Here, $\Xi_1 = \sqrt{\kappa q}x + \beta Jm$. We list the eigenvalues, their degeneracies and eigenvectors.

$$\begin{aligned}\lambda_1^{(1)\pm} &= \frac{1}{2}\{X \pm \sqrt{Y^2+Z}\}, \quad \text{degeneracy: 1 for each,}\\ X &= A + (n-1)B + P + 2(n-2)Q + \frac{(n-2)(n-3)}{2}R,\\ Y &= A + (n-1)B - P - 2(n-2)Q - \frac{(n-2)(n-3)}{2}R,\\ Z &= 2(n-1)\{2C + (n-2)D\}^2,\\ \epsilon_1^\alpha &= a, \epsilon_2^\alpha = 0, \epsilon_3^\alpha = 0, \eta^{\alpha\beta} = b,\end{aligned} \tag{A.14}$$

$$\begin{aligned}\lambda_1^{(2)} &= A + (n-1)B, \quad \text{degeneracy: 2,}\\ \epsilon_1^\alpha &= 0, \epsilon_\mu^\alpha = a', \eta^{\alpha\beta} = 0, \qquad \mu = 2 \text{ or } 3\ ,\end{aligned} \tag{A.15}$$

$$\lambda_2^{(1)\pm} = \frac{1}{2}\{X' \pm \sqrt{(Y')^2 + Z'}\}, \quad \text{degeneracy: } (n-1) \text{ for each}, \tag{A.16}$$
$$X' = A - B + P + (n-4)Q - (n-3)R,$$
$$Y' = A - B - P - (n-4)Q + (n-3)R,$$
$$Z' = 4(n-2)(C-D)^2,$$
$$\epsilon_1^\theta = c_1,\ \epsilon_1^\alpha = d_1,\ \epsilon_2^\beta = 0,\ \epsilon_3^\beta = 0,\ \alpha \neq \theta,\ \theta \text{ is some replica index},$$
$$\eta^{\theta\beta} = \eta^{\alpha\theta} = f, \eta^{\alpha\beta} = g, \qquad \alpha \neq \theta \text{ and } \beta \neq \theta,$$
$$\lambda_2^{(2)} = A - B, \quad \text{degeneracy: } 2(n-1), \tag{A.17}$$
$$\epsilon_\mu^\theta = c_2,\ \epsilon_\mu^\alpha = d_2,\ \epsilon_\nu^\beta = 0,\ \alpha \neq \theta,\ \nu \neq \mu,\ \theta \text{ is some replica index },\ \mu = 2 \text{ or } 3,$$
$$\eta^{\alpha\beta} = 0, \quad \text{for any } \alpha, \beta,$$
$$\lambda_3 = P - 2Q + R, \quad \text{degeneracy: } \tfrac{n(n-1)}{2} - n, \tag{A.18}$$
$$\epsilon_1^{\theta_1} = \epsilon_1^{\theta_2} = h_1, \epsilon_1^\alpha = h_2, \epsilon_\nu^\beta = 0, \theta_1 \neq \theta_2, \alpha \neq \theta_1, \alpha \neq \theta_2, \nu \neq 1,$$
$$\theta_1 \text{ and } \theta_2 \text{ are some replica indexes},$$
$$\eta^{\theta_1\theta_2} = u, \eta^{\theta_1\alpha} = \eta^{\theta_2\alpha} = v, \eta^{\alpha\beta} = w,$$
$$\alpha \neq \theta_1, \alpha \neq \theta_2 \text{ and } \beta \neq \theta_1, \beta \neq \theta_2. \tag{A.19}$$

Appendix A.2. Mixed state with 3 patterns

We put $m_1 = m_2 = m_3 = m$. Ξ_m is defined by

$$\Xi = \sqrt{\kappa q}x + \beta J \sum_\nu m_\nu \xi_\nu = \sqrt{\kappa q}x + \beta J m(\xi_1 + \xi_2 + \xi_3) \equiv \Xi_m. \tag{A.20}$$

Now, we list the non-zero elements of $\mathcal{G}$.

$$\mathcal{G}_{(\alpha\nu)(\alpha\nu)} \equiv \frac{\partial^2 G}{\partial m_\nu^{\alpha 2}} = -\beta J + (\beta J)^2\Big\{1 - [(\overline{\tanh \Xi_m})^2]\Big\} = A_1. \tag{A.21}$$

Since this quantity does not depend on ν, we define it A_1. In the below, we assume $\mu \neq \nu$ and α, β, γ and δ are all different.

$$\begin{aligned}\mathcal{G}_{(\alpha\nu)(\alpha\mu)} &\equiv \frac{\partial^2 G}{\partial m_\nu^\alpha \partial m_\mu^\alpha} = (\beta J)^2\Big\{[\langle \sigma^\alpha \sigma^\beta \xi^\nu \xi^\mu \rangle] - [\langle \sigma^\alpha \xi^\nu \rangle \langle \sigma^\alpha \xi^\mu \rangle]\Big\} \\ &= (\beta J)^2\Big\{[\xi^\nu \xi^\mu] - [\langle \sigma^\alpha \rangle^2 \xi^\mu \xi^\nu]\Big\} \\ &= -(\beta J)^2\Big[(\overline{\tanh \Xi_m})^2 \xi^\mu \xi^\nu\Big] = A_2, \end{aligned} \tag{A.22}$$

$$\begin{aligned}\mathcal{G}_{(\alpha\nu)(\beta\nu)} &\equiv \frac{\partial^2 G}{\partial m_\nu^\alpha \partial m_\nu^\beta} = (\beta J)^2\Big\{[\langle \sigma^\alpha \sigma^\beta \rangle] - [\langle \sigma^\alpha \rangle \xi^\nu \langle \sigma^\beta \rangle \xi^\nu]\Big\} \\ &= (\beta J)^2\Big\{[\overline{\tanh^2 \Xi_m}] - [(\overline{\tanh \Xi_m})^2]\Big\} \\ &= (\beta J)^2\Big\{q - [(\overline{\tanh \Xi_m})^2]\Big\} = B_1. \end{aligned} \tag{A.23}$$

At the last line, we used the following relations,

$$q = \Big[\langle \sigma^\alpha \sigma^\beta \rangle\Big] = \Big[\ \overline{\tanh^2 \Xi_m}\ \Big], \tag{A.24}$$
$$m_\nu^\alpha = \Big[\langle \sigma^\alpha \rangle \xi^\nu\Big] = \Big[\overline{\tanh \Xi_m} \xi^\nu\Big] = m_\nu = m. \tag{A.25}$$

$$\mathcal{G}_{(\alpha\nu)(\beta\mu)} \equiv \frac{\partial^2 G}{\partial m_\nu^\alpha \partial m_\mu^\beta} = (\beta J)^2 \Big\{ [\langle \sigma^\alpha \sigma^\beta \rangle \xi^\nu \xi^\mu] - [\langle \sigma^\alpha \rangle \langle \sigma^\beta \rangle \xi^\nu \xi^\mu] \Big\}$$
$$= (\beta J)^2 \Big\{ \overline{\tanh^2 \Xi_m} \xi^\nu \xi^\mu - [(\overline{\tanh \Xi_m})^2 \xi^\mu \xi^\nu] \Big\} = B_2, \tag{A.26}$$
$$\mathcal{G}_{(\alpha\beta)(\alpha\nu)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta} \partial m_\nu^\alpha} = \kappa\beta J \Big[\langle \sigma^\alpha \sigma^\beta \sigma^\alpha \xi^\nu \rangle - \langle \sigma^\alpha \sigma^\beta \rangle \langle \sigma^\alpha \xi^\nu \rangle \Big]$$
$$= \kappa\beta J \Big\{ \Big[\overline{\tanh \Xi_m} \xi^\nu \Big] - [\overline{\tanh^2 \Xi_m} \cdot \overline{\tanh \Xi_m} \xi^\nu] \Big\}$$
$$= \kappa\beta J \Big\{ m - [\overline{\tanh^2 \Xi_m} \cdot \overline{\tanh \Xi_m} \xi^\nu] \Big\} \equiv C, \tag{A.27}$$
$$\mathcal{G}_{(\alpha\beta)(\nu\gamma)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta} \partial m_\nu^\gamma} = \kappa\beta J \Big[\langle \sigma^\alpha \sigma^\beta \sigma^\gamma \rangle \xi^\nu - \langle \sigma^\alpha \sigma^\beta \rangle \langle \sigma^\gamma \rangle \xi^\nu \Big]$$
$$= \kappa\beta J \Big\{ [\overline{\tanh^3 \Xi_m} \xi^\nu] - [\overline{\tanh^2 \Xi_m} \cdot \overline{\tanh \Xi_m} \xi^\nu] \Big\} \equiv D, \tag{A.28}$$
$$\mathcal{G}_{(\alpha\beta)(\alpha\beta)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta}^2} = -\kappa + \kappa^2 \Big\{ 1 - [\langle \sigma^\alpha \sigma^\beta \rangle^2] \Big\}$$
$$= -\kappa + \kappa^2 \Big\{ 1 - [(\overline{\tanh^2 \Xi_m})^2] \Big\} \equiv P, \tag{A.29}$$
$$\mathcal{G}_{(\alpha\beta)(\alpha\gamma)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta} \partial q_{\alpha\gamma}} = \kappa^2 [\langle \sigma^\alpha \sigma^\beta \sigma^\alpha \sigma^\gamma \rangle - \langle \sigma^\alpha \sigma^\beta \rangle^2]$$
$$= \kappa^2 \Big\{ [\overline{\tanh^2 \Xi_m}] - [(\overline{\tanh^2 \Xi_m})^2] \Big\}$$
$$= \kappa^2 \Big\{ q - [(\overline{\tanh^2 \Xi_m})^2] \Big\} \equiv Q, \tag{A.30}$$
$$\mathcal{G}_{(\alpha\beta)(\gamma\delta)} \equiv \frac{\partial^2 G}{\partial q_{\alpha\beta} \partial q_{\gamma\delta}} = \kappa^2 [\langle \sigma^\alpha \sigma^\beta \sigma^\gamma \sigma^\delta \rangle - \langle \sigma^\alpha \sigma^\beta \rangle^2]$$
$$= \kappa^2 \Big\{ [\overline{\tanh^4 \Xi_m}] - [(\overline{\tanh^2 \Xi_m})^2] \Big\} \equiv R. \tag{A.31}$$

We list the eigenvalues, their degeneracies and eigenvectors.

$$\lambda_1^{(1)\pm} = \frac{1}{2} \{ X \pm \sqrt{Y^2 + Z} \}, \quad \text{degeneracy: 1 for each,} \tag{A.32}$$
$$X = A + (n-1)B + P + 2(n-2)Q + \frac{(n-2)(n-3)}{2} R,$$
$$Y = A + (n-1)B - P - 2(n-2)Q - \frac{(n-2)(n-3)}{2} R,$$
$$Z = 6(n-1)\{2C + (n-2)D\}^2,$$
$$A = A_1 + (n-1)B_1, \ B = \frac{2(A_2 + (n-1)B_2)}{n-1},$$
$$\epsilon_1^\alpha = \epsilon_2^\alpha = \epsilon_3^\alpha = a, \eta^{\alpha\beta} = b,$$
$$\lambda_1^{(2)} = A_1 - A_2 + (n-1)(B_1 - B_2), \quad \text{degeneracy: 2,} \tag{A.33}$$
$$\epsilon_1^\alpha = a_1', \epsilon_2^\alpha = a_2', \epsilon_3^\alpha = a_3', \ a_1' + a_2' + a_3' = 0, \ \eta^{\alpha\beta} = 0,$$
$$\lambda_2^{(1)\pm} = \frac{1}{2} \{ X' \pm \sqrt{(Y')^2 + Z'} \}, \quad \text{degeneracy: } (n-1) \text{ for each,} \tag{A.34}$$
$$X' = A_1 - B_1 + 2(A_2 - B_2) + P + (n-4)Q - (n-3)R,$$
$$Y' = A_1 - B_1 + 2(A_2 - B_2) - P - (n-4)Q + (n-3)R,$$
$$Z' = 12(n-2)(C-D)^2,$$

$$
\begin{aligned}
& \epsilon_1^\theta = c_1,\ \epsilon_1^\alpha = d_1, \epsilon_2^\theta = c_2,\ \epsilon_2^\alpha = d_2, \epsilon_3^\theta = c_3,\ \epsilon_3^\alpha = d_3, \\
& \qquad \alpha \neq \theta,\ \theta \text{ is some replica index}, \\
& \eta^{\theta\beta} = \eta^{\alpha\theta} =, \eta^{\alpha\beta} = g, \qquad \alpha \neq \theta \text{ and } \beta \neq \theta, \\
\lambda_2^{(2)} \quad & = A_1 - A_2 - (B_1 - B_2), \quad \text{degeneracy: } 2(n-1), \qquad \text{(A.35)} \\
& \epsilon_1^\theta = c_1,\ \epsilon_1^\alpha = d_1, \epsilon_2^\theta = c_2,\ \epsilon_2^\alpha = d_2, \epsilon_3^\theta = c_3,\ \epsilon_3^\alpha = d_3, \\
& \qquad \alpha \neq \theta,\ \theta \text{ is some replica index}, \\
& \eta^{\alpha\beta} = 0, \qquad \text{for any } \alpha, \beta, \\
& \eta^{\alpha\beta} = 0, \quad \text{for any } \alpha, \beta, \\
& c_i = (1-n)d_i,\ d_1 + d_2 + d_3 = 0, \\
\lambda_3 \quad & = P - 2Q + R, \quad \text{degeneracy: } \tfrac{n(n-1)}{2} - n, \qquad \text{(A.36)} \\
& \epsilon_\mu^{\theta_1} = \epsilon_\mu^{\theta_2} = r_\mu,\ \epsilon_\mu^\alpha = s_\mu, \quad \mu = 1,2,3, \theta_1 \neq \theta_2,\ \alpha \neq \theta_1, \alpha \neq \theta_2, \\
& \ \theta_1 \text{ and } \theta_2 \text{ are some replica indexes}, \\
& \eta^{\theta_1\theta_2} = u,\ \eta^{\theta_1\alpha} = \eta^{\theta_2\alpha} = v,\ \eta^{\alpha\beta} = w, \\
& \qquad \alpha \neq \theta_1, \alpha \neq \theta_2 \text{ and } \beta \neq \theta_1, \beta \neq \theta_2.
\end{aligned}
$$

Appendix A.3. Spin Glass

For the spin glass solution, $m_\mu = 0$ and $q \neq 0$. Thus, we only have to put $m = 0, C = D = 0$ in the quantities for the Hopfield attractor. Thus, we obtain the following eigenvalues:

$$\lambda_1^{(1)+} = A + (n-1)B = \lambda_1^{(2)} = \lambda_1^{(3)}, \qquad \text{(A.37)}$$

$$\lambda_1^{(1)-} = P + 2(n-2)Q + \frac{(n-2)(n-3)}{2}R, \qquad \text{(A.38)}$$

$$\lambda_2^{(1)+} = A - B, \qquad \text{(A.39)}$$

$$\lambda_2^{(1)-} = P + (n-1)Q - (n-3)R = \lambda_2^{(2)} = \lambda_2^{(3)}, \qquad \text{(A.40)}$$

$$\lambda_3 \quad = P - 2Q + R. \qquad \text{(A.41)}$$

References

[1] Coolen A C C, Penney R W and Sherrington D 1993 *Phys. Rev.* B **48** 16116
[2] Penney R W, Coolen A C C and Sherrington D 1993 *J. Phys. A: Math. Gen.* **26** 3681
[3] Penney R W and Sherrington D 1994 *J. Phys. A: Math. Gen.* **27** 4027
[4] Dotsenko V, Franz S and Mézard M 1994 *J. Phys. A: Math. Gen.* **27** 2351
[5] Uezu T and Coolen A C C 2002 *J. Phys. A: Math. Gen.* **35** 2761
[6] Abe K 2008 Masters thesis, Graduate School of Humanities and Sciences, Nara Women's University, Nara [In Japanese]
[7] Hara K, Miyoshi S, Uezu T and Okada M 2008 *Proceeding of International Conference on Neural Information Processing* .
[8] Kimoto T et al 2008 in preparation
[9] de Almeida J R L and Thouless D J 1978 *J. Phys. A: Math. Gen.* **11** 983